

Schwippert, Knut

Daten für die Schulentwicklung – auf die Perspektive kommt es an.

Interpretationen und Fehlinterpretationen von Daten aus der Schulinspektion

Pietsch, Marcus [Hrsg.]; Scholand, Barbara [Hrsg.]; Schulte, Klaudia [Hrsg.]: Schulinspektion in Hamburg. Der erste Zyklus 2007 - 2013: Grundlagen, Befunde und Perspektiven. Münster [u.a.] : Waxmann 2015, S. 157-176. - (HANSE - Hamburger Schriften zur Qualität im Bildungswesen; 15)



Quellenangabe/ Reference:

Schwippert, Knut: Daten für die Schulentwicklung – auf die Perspektive kommt es an. Interpretationen und Fehlinterpretationen von Daten aus der Schulinspektion - In: Pietsch, Marcus [Hrsg.]; Scholand, Barbara [Hrsg.]; Schulte, Klaudia [Hrsg.]: Schulinspektion in Hamburg. Der erste Zyklus 2007 - 2013: Grundlagen, Befunde und Perspektiven. Münster [u.a.] : Waxmann 2015, S. 157-176 - URN: urn:nbn:de:0111-pedocs-115338 - DOI: 10.25656/01:11533

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-115338>

<https://doi.org/10.25656/01:11533>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Marcus Pietsch, Barbara Scholand,
Klaudia Schulte (Hrsg.)

Schulinspektion in Hamburg

Der erste Zyklus 2007–2013:
Grundlagen, Befunde,
Perspektiven



HANSE
Hamburger Schriften zur Qualität
im Bildungswesen

herausgegeben von
Detlef Fickermann und Knut Schwippert

Band 15



Waxmann 2015
Münster • New York

Marcus Pietsch, Barbara Scholand,
Klaudia Schulte (Hrsg.)

Schulinspektion in Hamburg

Der erste Zyklus 2007–2013:
Grundlagen, Befunde und Perspektiven



Das gedruckte Buch finden Sie [hier](#).



Waxmann 2015
Münster • New York

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten
sind im Internet über <http://dnb.d-nb.de> abrufbar.

HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Band 15

ISSN 1864-2225

ISBN 978-3-8309-3278-9

© Waxmann Verlag GmbH, 2015
Steinfurter Straße 555, 48159 Münster

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Pleßmann Design, Ascheberg
Umschlagfoto: © Robert Kneschke – Fotolia.de
Lektorat und Satz: Judith Zimmer, Hamburg
Druck: Mediaprint, Paderborn

Gedruckt auf alterungsbeständigem Papier,
säurefrei gemäß ISO 9706

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.
Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages
in irgendeiner Form reproduziert oder unter Verwendung elektronischer
Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Inhalt

Vorwort.	9
---------------	---

Teil 1: Die Grundlagen der Schulinspektion

Norbert Maritzen

Schulinspektion – Aspekte der Transformation von Governance-Strukturen . . .	13
--	----

Norbert Maritzen & Herbert Altrichter

Schulinspektion in Hamburg – ein Gespräch über Startbedingungen und Entwicklungen	37
--	----

Martina Diedrich

Aufbau und Rolle der Schulinspektion Hamburg	57
--	----

Thomas Brüsemeister

Erfahrungsbereiche der Hamburger Schulinspektion – Governance-analytische Programmtheorie	77
--	----

Susanne Ulrich

Das Qualitätsleitbild der Schulinspektion Hamburg	101
---	-----

Teil 2: Die Arbeit der Schulinspektion

Marcus Pietsch, Ann-Katrin van den Ham & Olaf Köller

Wirkungen von Schulinspektion: Ein Rahmen zur theoriegeleiteten Analyse von Schulinspektionseffekten	117
---	-----

Moritz G. Sowada

Expertenurteile – Achillesferse oder Trumpf der Schulinspektion?	137
--	-----

Knut Schwippert

Daten für die Schulentwicklung – auf die Perspektive kommt es an	157
--	-----

<i>Barbara Scholand unter Mitarbeit von Stephanie Graw</i>	
Wie Schulinspektor/innen fragen und Grundschüler/innen antworten	177

<i>Peter Schulze & Christiane von Schachtmeyer</i>	
Der Blick aus der Praxis – die Schulformexpertinnen und -experten	219

Teil 3: Die Befunde der Schulinspektion

<i>Melanie Ehren & Jaap Scheerens</i>	
Evidenzbasierte Referenzrahmen zur Schulqualität als Grundlage von Schulinspektion	233

<i>Holger Gärtner</i>	
Zusammenhang von Schul- und Unterrichtsqualität und schulischen Rahmenbedingungen	273

<i>Marcus Pietsch & Timo Ehmke</i>	
Die Qualität des Unterrichts an SINUS-Grundschulen. Eine Programmevaluation mithilfe von Daten der Schulinspektion Hamburg	295

<i>Klaudia Schulte & Markus Lücken</i>	
Der Einfluss schulischer Prozesse auf die Lernentwicklung der Schülerinnen und Schüler an weiterführenden Schulen in Hamburg	317

<i>Susanne Peters</i>	
Eltern als Stakeholder von Schule. Erkenntnisse über die Sichtweise von Eltern durch die Hamburger Schulinspektion.	341

Teil 4: Die Zukunft der Schulinspektion

<i>Sabine Müller</i>	
Zukunft der externen Evaluation – Trends und Perspektiven in Deutschland. .	369

<i>Hans-Jürgen Kuhn</i>	
Rechtliche, strukturelle und politische Rahmenbedingungen für externe Evaluation in Deutschland.	385

Martina Diedrich

Der zweite Zyklus der Schulinspektion Hamburg: ein Ausblick. 419

*Martin Spiewak, Angela Hellberg, Claudia Lemke, Gerrit Petrich,
Norbert Rosenboom, Mina Zulal*

Die Perspektive der Stakeholder: Ein Interview von Martin Spiewak 437

Autorinnen und Autoren 455

Vorwort

Vor rund zehn Jahren wurden die ersten Schulinspektorate in Deutschland eingeführt. Als Teil einer veränderten Steuerungsstrategie, die die Qualitätssicherung und -entwicklung von Schulen ins Zentrum setzt, sollen sie die Qualität schulischer Prozesse evaluieren, prozessuale Mindeststandards in Schule und Unterricht sichern und darüber hinaus Schulentwicklung stimulieren.

Die Hamburger Schulinspektion ist keine Inspektion der allerersten Stunde, wurde sie doch erst im Jahr 2007 offiziell eingerichtet. Zu diesem Zeitpunkt waren in vielen anderen Bundesländern schulische Evaluatorinnen und Evaluatoren bereits zwei bis drei Jahre im Feld und hatten die unterschiedlichsten Verfahren entwickelt und verschiedene Konzepte erprobt. Von ihren vielfältigen Erfahrungen mit diesem in Deutschland dahin unbekannten Instrument konnte die Hamburger Schulinspektion profitieren.

Gleichwohl hatte man in Hamburg bereits seit der Jahrtausendwende erste Konzepte zur externen Prozessevaluation an Einzelschulen in Hamburg diskutiert und entwickelt. Den Impuls hierfür gaben insbesondere die Ergebnisse der Längsschnittstudie LAU (Aspekte der Lernausgangslage und der Lernentwicklung), die seit Mitte der 1990er Jahre regelmäßig durchgeführt wurde. Es vergingen jedoch noch mehrere Jahre, bis die erste Hamburger Schule tatsächlich durch ein Inspektionssteam evaluiert wurde. Seitdem bewegt sich die Schulinspektion im Hamburger Schulsystem stets im Spannungsfeld von Administration, Praxis und Wissenschaft und muss dabei ihren eigenen sowie den Ansprüchen der verschiedenen schulischen und außerschulischen Stakeholder gerecht werden.

In ihrem ersten Zyklus, der im Jahr 2007 begann und bis zum Jahr 2013 andauerte, inspizierte die Schulinspektion 375 Hamburger Schulen. Während dieser sechs Jahre nahmen etwa 10.000 Lehrerinnen und Lehrer, gut 36.000 Eltern, rund 92.000 Schülerinnen und Schüler sowie etwa 2.000 Vertreterinnen und Vertreter Hamburger Unternehmen an den schriftlichen Befragungen der Inspektion teil. Darüber hinaus wurden fast 30.000 Unterrichtssequenzen durch die Inspektorinnen und Inspektoren beobachtet und etwa 2.000 Interviews mit Schulleitungen, Eltern- und Schülerräten sowie Lehrerkollegien und weiteren Mitarbeiterinnen und Mitarbeitern der inspizierten Schulen, aber auch mit Vertreterinnen und Vertretern von Ausbildungsbetrieben geführt.

Die Stärke des vorliegenden Bandes ist es, so meinen wir, dass er der besonderen Stellung der Schulinspektion im Gesamtgefüge des Hamburger Bildungssystems Rechnung trägt, indem er Akteure aus den verschiedenen Bereichen und mit unterschiedlichen Perspektiven zu Wort kommen lässt. Ein Teil der Beiträge ist empirisch ausgerichtet und analysiert die Arbeit der Inspektion oder nutzt die im Rahmen der Schulinspektion Hamburg erhobenen Daten für weiterführende

Analysen zum Hamburger Schulsystem. Andere Beiträge wiederum setzen sich mit den Rahmenbedingungen, Grundlagen und Folgen von Schulinspektion auseinander oder widmen sich den Erfahrungen und Erwartungen einzelner Akteure und Akteursgruppen.

Wir bedanken uns herzlich bei allen Autorinnen und Autoren dieses Bandes für die kooperative, kollegiale und anregende Zusammenarbeit. Insbesondere haben wir uns darüber gefreut, dass die vielen Kolleginnen und Kollegen aus diversen Kontexten und Arbeitsfeldern umstandslos bereit waren, ihre Expertise in Form fundierter und kritischer Beiträge sowie Gutachten beizusteuern. Sehr glücklich sind wir vor allem auch darüber, dass sich die Vertreterinnen und Vertreter der schulischen Stakeholder aktiv in diese Publikation eingebracht haben.

In der Zusammenarbeit mit all diesen verschiedenen Personen ist uns deutlich geworden, dass es institutionen- und akteursübergreifend ein großes Interesse daran gibt, über gute Bildung in und für Hamburg ins Gespräch zu kommen und sie gemeinsam zu gestalten. Dafür bieten – nicht nur, aber auch – die Schulinspektion und ihre Ergebnisse konstruktive Gesprächsanlässe. In diesem Sinne wünschen wir Ihnen eine ertragreiche Lektüre!

Hamburg im Juni 2015

Marcus Pietsch, Barbara Scholand & Klaudia Schulte

Daten für die Schulentwicklung – auf die Perspektive kommt es an

Interpretationen und Fehlinterpretationen von Daten aus der Schulinspektion

Knut Schwippert

Zusammenfassung

Dem Schulinspektionsbericht von 2009/2010 ist zu entnehmen, dass neben der Aufgabe, den Einzelschulen Rückmeldung über den Stand ihrer Qualitätsentwicklung zu geben, die Schulinspektion auch den Auftrag zum Bildungsmonitoring hat. Das damit betraute Institut stellt Daten über die Qualität des Hamburger Schulsystems zur Verfügung, die vor dem Hintergrund unterschiedlicher Fragestellungen ausgewertet werden können. Neben dem Impuls, direkte Informationen an Schulen und für die Unterrichtsentwicklung vor Ort bereitzustellen, können die Daten auch im Sinne eines Bildungsmonitorings genutzt werden. Werden Bedingungen, unter denen die Daten erhoben wurden, bei der rechnerischen Auswertung nicht berücksichtigt, kann es jedoch zu Fehlinterpretationen kommen. Cronbach hat schon 1976 die Auffassung vertreten, dass die (bis dato) durchgeführte Unterrichtsforschung mehr verschleiert, denn zum Vorschein gebracht habe. Dieser Beitrag möchte für die jeweiligen Tragweiten unterschiedlicher Analyseansätze sensibilisieren und dies anhand der Schulinspektionsdaten durch vergleichend gegenübergestellte Mehrebenenanalysen verdeutlichen.

Dem Schulinspektionsbericht von 2009/2010 ist zu entnehmen (Institut für Bildungsmonitoring 2011, S. 6), dass neben der Aufgabe, den Einzelschulen Rückmeldung über den Stand ihrer Qualitätsentwicklung zu geben, die Schulinspektion auch den Auftrag zum Bildungsmonitoring hat. Hierbei stellt sie übergreifende Erkenntnisse über die Qualität des Hamburger Schulsystems im Rahmen ihrer Jahresberichte zur Verfügung. Die im Rahmen der Inspektionen erhobenen Daten können dann vor dem Hintergrund ganz unterschiedlicher Fragestellungen ausgewertet werden. Neben dem Impuls, direkte Informationen an Schulen und für die Unterrichtsentwicklung vor Ort bereitzustellen, können diese auch im Sinne eines Bildungsmonitorings für Fragen, die für die Bildungsadministration bzw. behördliche Bildungsplanung relevant sind, ausgewertet werden. Und schließlich bieten die Daten auch eine mögliche Basis zur Beantwortung wissenschaftlicher Frage-

stellungen, die nicht unmittelbar Wirkungen im Unterricht bzw. in der Bildungssteuerung entfalten. Je nachdem, welche Fragestellungen mit den Inspektionsdaten beantwortet werden sollen, ist sicher zu stellen, dass der besonderen Struktur der Daten bei den empirischen Analysen Rechnung getragen wird. So sind bei Analysen einzelner Klassen bzw. Schulen ggf. andere Verfahren anzuwenden als bei solchen, die das gesamte Bildungssystem in den Blick nehmen. Dabei ist die Zuordnung von Fragestellung, der Identifizierung der angemessenen Analyseebene (Individuum, Klasse, Schule, Bezirk, System) und der jeweils angemessenen Analysemethode nicht beliebig austauschbar. Dieser Beitrag möchte anhand von Re-Analysen der Hamburger Schulinspektionsdaten für die jeweiligen Tragweiten unterschiedlicher Analyseansätze sensibilisieren. Auch wenn dieser Beitrag insbesondere mit der Perspektive einer wissenschaftlich orientierten Sekundäranalyse geschrieben wurde, so sind die methodologischen Anmerkungen auch für weiterführende Analysen im Rahmen der Hamburger Schulinspektion sowie für Analysen anderer externer Evaluationen in Deutschland relevant – wenn spezifische Perspektiven für die Auswertung eingenommen werden.

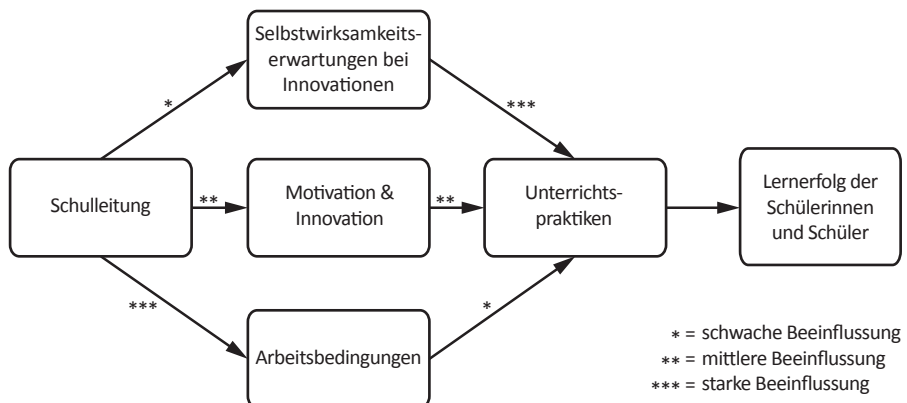
Aktuelle Untersuchungen zur Nutzung von Daten aus der Schulinspektion können auf eine Reihe von Daten zurückgreifen, die eine empirische Absicherung vermuteter bzw. in Einzelfällen beobachteter Zusammenhänge erlauben. Daten aus Befragungen und Unterrichtsbeobachtungen im Rahmen von Schulinspektionen basieren auf einem komplexen Stichproben- und Testdesign (vgl. Adams & Wu 2000; Institut für Bildungsmonitoring 2009, 2011; Institut für Bildungsmonitoring und Qualitätsentwicklung 2012; Martin et al. 2007; Martin et al. 1999). Werden Bedingungen, unter denen die Daten erhoben wurden, bei der rechnerischen Auswertung nicht berücksichtigt, kann es zu Fehlinterpretationen kommen. Cronbach hat schon 1976 in seiner Veröffentlichung *Research on Classrooms and Schools* (Cronbach 1976) die Auffassung vertreten, dass die (bis dato) durchgeführte Unterrichtsforschung mehr verschleiert, denn zum Vorschein gebracht habe.

Cronbachs Kritik richtet sich insbesondere darauf, dass bei Analysen von Daten, in denen Schülerinnen und Schüler ganzer Klassen bzw. Schulen getestet wurden, nicht angemessen berücksichtigt würde, dass sich Schüler einer Schulklasse (oder einer Schule) untereinander ähnlicher sind als die gleiche Anzahl von Schülern, die zufällig aus verschiedenen Klassen stammen. Um eine angemessene Berücksichtigung dieser sogenannten Klumpungseffekte geht es in diesem Beitrag (vgl. z. B. Cochran 1972; Kish 1995). Aber auch wenn die Daten auf zu hoher Ebene zusammengefasst werden, kann sich ein Problem ergeben: Bei der Betrachtung von solchen aggregierten Daten kann es zu einer Übergeneralisierung kommen, wenn die Ergebnisse gleichförmig auf alle beteiligten Bildungseinheiten, also z. B. Schulen oder Klassen, übertragen werden. Solche Form der Interpretation (im Sinne eines: ‚Eigentlich sind wir alle gleich.‘) kann dazu führen, dass problematische, aber auch vorbildliche Situationen übersehen werden. Der Kern der zuvor skizzierten Kritik

ist, dass es bei Untersuchungen im Bereich der Schul- und Unterrichtsforschung in Abhängigkeit von der jeweils zu beantwortenden Fragestellung zur fehlerhaften Berechnung z. B. von Schul- bzw. Unterrichtseffekten kommen kann, wenn die Gruppierung von Individuen in Gruppen unberücksichtigt bleibt (vgl. Ditton 1998; Janke 2006; Schwippert 2002) und die zu beantwortende Frage nicht mit der verwendeten Analyseverfahren korrespondiert. Dies hieße für die Schulinspektion, dass dann, wenn generelle Effekte auf Bildungssystemebene zu untersuchen sind, auch diese Ebene in den Blick genommen werden muss. Wenn es aber um das Wirken oder das Zusammenspiel von Qualitätsmerkmalen zwischen Schulen geht, müssen die Beobachtungen auch für die Beschreibung von Unterschieden zwischen den Schulen geeignet sein. Schließlich muss es analog gelten, Qualitätsmerkmale im Unterricht zu erheben, wenn Unterschiede des Unterrichts zwischen verschiedenen Klassen betrachtet werden sollen. Im letzten Fall wäre eine differenzielle Diagnostik bzw. ein differenzieller Blick auf verschiedene Merkmale des Unterrichts zu richten, während bei der Beurteilung der Bildungssystem- bzw. Schulqualität eine umfassendere und für die einzelnen Schulen bzw. Unterrichte spezifischere Diagnostik (die dann über die verschiedenen Bildungseinrichtungen hinweg ihre Breite belegen) methodisch angemessen ist.

Für die nachfolgend präsentierten Betrachtungen auf verschiedenen Ebenen des Bildungssystems wird hier auf das Modell (siehe Abbildung 1) des Schulleitungshandelns von Leithwood et al. (2006) zurückgegriffen (zitiert nach Pietsch 2014).

Abbildung 1: Modell des Einflusses von Schulleitungen auf die Unterrichtsgestaltung von Lehrkräften (adaptierte Übersetzung von Pietsch 2014)

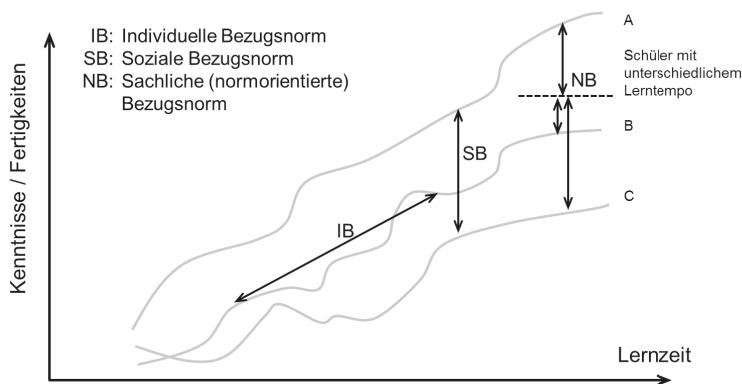


In diesem Modell ist die empirisch abgesicherte Wirkungskette des Schulleitungshandelns auf „Selbstwirksamkeitserwartungen bei Innovationen“, auf „Motivation und Identifikation“ und auf „Arbeitsbedingungen“ und deren Wirkungen auf die

„Unterrichtspraktiken“, die wiederum direkte Effekte auf den Lernerfolg von Schülerinnen und Schülern haben, abgebildet. Aus diesem Wirkungsmodell sei hier der Pfad (die Wirkung) des Schulleitungshandelns auf die „Motivation und Identifikation“ von Lehrkräften herausgegriffen. Anhand des gewählten Beispiels wird verdeutlicht, zu welchen unterschiedlichen Einschätzungen man bezüglich des Zusammenhangs zwischen der Zufriedenheit der Lehrkräfte und dem Schulleitungshandeln in Bezug auf den Aspekt der Führung kommt, wenn mit Daten auf unterschiedlichen Ebenen des Bildungssystems argumentiert wird.

Doch warum sollten bei der Analyse des Zusammenhangs des Führungshandelns von Schulleitungen mit der Zufriedenheit der Lehrkräfte unterschiedliche Ebenen des Bildungssystems betrachtet werden? Ein Blick auf drei mögliche Bezugsnormorientierungen bei Leistungsbewertungen von Schülerinnen und Schülern durch Lehrkräfte soll als Analogiebeispiel zur Beantwortung dieser Frage herangezogen werden. Nach Rheinberg (2001) können individuelle, soziale und sachliche Beurteilungsmaßstäbe unterschieden werden. (Alternativ werden diese auch intraindividuelle, interindividuelle und kriteriale Bezugsnormen genannt, vgl. Krapp & Weidenmann 2001.) Jede dieser Bewertungsnormen hat ihre spezifischen Vor- und Nachteile (Rheinberg bezeichnet Letztere als „blinde Flecke“), sodass keine der einzelnen Normen pauschal als besser oder weniger geeignet bewertet werden kann.

Abbildung 2: Bezugsnormorientierungen nach Rheinberg (2001, S. 62) (eigene Darstellung)



In Abbildung 2 sind die (fiktiven) Lernverläufe von drei sich unterschiedlich entwickelnden Kindern abgetragen. Bei der Bewertung des Lernstandes lässt sich zunächst die *individuelle Bezugsnormorientierung* (IB) anlegen, die hier

die Veränderungen einer einzelnen Person über die Zeit in den Blick nimmt und hierbei – orientiert an einem Ausgangspunkt – den Lernzuwachs bewertet. Bei der *sozialen Bezugsnormorientierung* (SB) wird der relative Kenntnisstand eines Kindes im Vergleich mit – in der Regel – den anderen Kindern in derselben Schulklasse bewertet. Diese Bezugsnormorientierung ist im deutschen Schulsystem eine sehr oft genutzte Bezugsnorm. Schließlich gibt es die *sachliche (normorientierte) Bezugsnorm* (NB), bei der, bezogen auf ein vorgegebenes Niveau, die Leistungen der Kinder beurteilt werden. Für Letzteres sind das Zentralabitur oder die Führerscheinprüfung Beispiele.

Der blinde Fleck bei der *individuellen Bezugsnormorientierung* ist, dass – trotz detailliertem Wissen über die individuellen Veränderungen von Schülerinnen und Schülern – keine Information vorliegt, ob der Schüler oder die Schülerin altersgemäß eher gut oder eher schwach erscheint, es also nicht beurteilt werden kann, ob besondere Förderung notwendig ist oder nicht. Der blinde Fleck der *sozialen Bezugsnormorientierung* ist die fehlende Sensibilität gegenüber der generellen Entwicklung der Gruppe. Anhand der *sozialen Bezugsnormorientierung* kann nicht festgestellt werden, ob sich die Gruppenmitglieder weiterentwickelt haben oder nicht. Hieraus ergibt sich der Nachteil, dass schwache Schülerinnen oder Schüler einer Lerngruppe, obwohl sie wie alle anderen in der Gruppe einen bedeutenden Lernzuwachs (beurteilt nach der *individuellen Bezugsnorm*) realisieren konnten, auch weiterhin als schwach bewertet werden. Die *sachliche (normorientierte) Bezugsnorm* hat insbesondere dann einen blinden Fleck, wenn als Bewertungsergebnis nur der Befund „bestanden“ oder „nicht bestanden“ zurückgemeldet wird. In beiden Fällen ist nicht deutlich, ob das geforderte Lernziel knapp oder deutlich erreicht bzw. verfehlt wurde.

Lehrkräfte wünschen sich in der Regel die Möglichkeit der *individuellen Bezugsnormorientierung*, sind aber gehalten, in der Regel die *soziale Bezugsnorm* anzuwenden. (Diese ist auch die, die die Schülerinnen und Schüler untereinander anlegen, wenn sie Noten oder Zeugnisse untereinander vergleichen.) Schließlich wäre die *sachliche (normorientierte) Bezugsnorm* der fairste Ansatz, wenn es um die Zuweisung von Bildungsressourcen bzw. -zertifikaten geht, da hierbei die zufällige Zusammensetzung der Gruppe, aus der eine Person kommt, keine Rolle spielt. Mit der Frage nach der Wahl der Bezugsnorm ist auch die Frage verbunden, was im Fokus der Bewertung steht (individuelle Entwicklung, Status in einer Gruppe, Normerreicherung).

Übertragen auf die Betrachtung der Daten aus der Schulinspektion hieße dies, zunächst die Frage zu klären, mit welcher Perspektive bzw. auf welcher Ebene die Informationen aus der Schulinspektion genutzt werden sollen. Ist die Frage genereller Zusammenhänge im Hamburger Schulsystem zu klären, steht eine spezifische Schulform im Fokus der Betrachtung oder sind die Einzelschulen oder sogar der Unterricht der Lehrkräfte bei der Inspektion im Zentrum der Betrachtung?

Im Folgenden wird gezeigt, dass es – bei gleichen Datenlagen – zu unterschiedlichen und differenzierten Befunden führen kann, wenn verschiedene Betrachtungsperspektiven eingenommen werden. Hierbei wird verdeutlicht, dass keine Perspektive „die richtige“ ist – vielmehr wird argumentiert, dass es auf die Fragestellung ankommt, welche Perspektive eingenommen werden sollte oder sogar muss.

1. Empirische Annäherung

1.1 Datenlage

Für die Analysen stehen die Daten der Hamburger Schulinspektion der ersten Inspektionswelle zur Verfügung. Hierbei wurden im Zeitraum von 2007 bis 2013 alle Hamburger Schulen einmal besucht und anhand der im Orientierungsrahmen für die Qualitätsentwicklung an Hamburger Schulen (Behörde für Bildung und Sport 2006) genannten Qualitätsaspekte inspiziert. Zur Konzeption, Organisation, Durchführung und Auswertung von Daten aus der Schulinspektion sei auf die seit der Einführung der Hamburger Schulinspektion regelmäßig erscheinenden Jahresberichte verwiesen (vgl. Institut für Bildungsmonitoring 2009, 2011; Institut für Bildungsmonitoring und Qualitätsentwicklung 2012; Beiträge von Diedrich in diesem Band).

Tabelle 1: Anzahl von befragten Lehrkräften der ersten Inspektionskohorte nach Schulform und Schuljahr

Schulform Schuljahr	Grund- schule	GHR- Schule	Gesamt- schule	Gym- nasium	Sonder-/ Förder- schule	beruf- liche Schule	Stadtteil- schule	Gesamt
2006/2007	99	116	13	27	0	26	0	281
2007/2008	364	194	246	426	38	369	0	1637
2008/2009	186	122	313	436	230	369	0	1656
2009/2010	445	82	241	485	164	448	0	1865
2010/2011	598	0	0	539	94	229	573	2033
2011/2012	610	0	0	224	21	168	152	1179
2012/2013	290	0	0	0	118	66	132	606
Gesamt	2592	514	813	2137	665	1675	857	9253

Um die Betrachtungen zu Effekten unterschiedlicher Gruppierungen von Daten nicht zu komplex und damit unübersichtlich zu gestalten, wird bei dem Vergleich zwischen Schulen ausschließlich auf die Schulinspektion der Grundschulen zurückgegriffen. Dies geschieht aus zwei Gründen: Die Schulform „Grundschule“ hat im Laufe des ersten Schulinspektionszyklus keine wesentlichen Strukturänderung

im Rahmen von Schulstrukturdebatten erfahren. Die Grundschule ist – da sie als Gesamtschulsystem angelegt ist – auch die zahlenmäßig stärkste Bildungseinrichtung mit $N=163$ Schulen und $N=2.592$ Lehrkräften. In Tabelle 1 ist die Verteilung der befragten Lehrkräfte nach Schulform und Schuljahr aufgeführt. In Tabelle 2 ist die Verteilung der inspizierten Schulen nach Schulform und Schuljahr eingetragen.

Tabelle 2: Anzahl von inspizierten Schulen der ersten Inspektionskohorte nach Schulform und Schuljahr

Schulform Schuljahr	Grund- schule	GHR- Schule	Gesamt- schule	Gym- nasium	Sonder-/ Förder- schule	beruf- liche Schule	Stadtteil- schule	Gesamt
2006/2007	8	5	1	1	0	1	0	16
2007/2008	24	9	7	14	1	10	0	65
2008/2009	14	7	8	13	11	10	0	63
2009/2010	31	4	7	14	4	13	0	73
2010/2011	32	0	0	18	4	6	14	74
2011/2012	39	0	0	6	1	4	5	55
2012/2013	15	0	0	0	2	1	2	20
Gesamt	163	25	23	66	23	45	21	366

1.2 Betrachtung der hierarchischen Datenstruktur

Wenn im Umgang mit Daten aus der Schulinspektion unberücksichtigt bleibt, dass Lehrkräfte in Schulen und diese wiederum in Schulformen organisiert sind, kann es zu fehlerhaften Schuleffektmessungen kommen (vgl. Treiber 1980). Aufgrund dieser sogenannten ‚geschachtelten‘ (hierarchischen) Datenstrukturen sind bei der Untersuchung z. B. der wahrgenommenen Führung durch die Schulleitungen feststellbare Unterschiede nicht nur zwischen den Lehrkräften, sondern auch zwischen den Schulen (Schulformen etc.) zu erwarten. Die Notwendigkeit einer getrennten Betrachtung von Innerhalb- (intra-) und Zwischen- (inter-) Schuleffekten belegt Treiber analog anhand der Betrachtung von Schülerinnen und Schülern in Klassen und in Schulen (vgl. ebd., S. 212). Seiner Auffassung nach sind „Analysen auf Schülerebene [...] nur dann korrekt, wenn die Gruppierung von Schülern in ihren Schulklassen ohne nennenswerten Einfluß“ auf das untersuchte Merkmal bleibt. Analoges gilt für die Gruppierung von Lehrkräften in Schulen. Überträgt man somit dieses Argument auf die Schulinspektion, gilt dies für die Gruppierung der Lehrkräfte (in Schulen) und damit möglicherweise auch für eine fehlerhafte Ermittlung der im Fokus stehenden (Schul-)Effekte. Bei den entsprechenden Analysen ist daher die Schule als Analyseeinheit zu berücksichtigen.

Der Effekt, dass sich Mitglieder einer realen Gruppe untereinander ähnlicher sind als eine gleich große, aber zufällig ausgewählte Gruppe von Personen, wird in der Empirie (Stichprobentheorie) Klumpungseffekt genannt. Er besagt, dass eine Gruppen von Personen, die gleiche Erfahrungen, gleiche Motivation und z. B. auch vergleichbare Bildungswege bzw. -erfahrungen haben, sich auch in anderen Merkmale (z.B. Interessen, Fähigkeiten, Fertigkeiten) ähnlicher sind als eine gleich große Gruppe von Personen, die entsprechende gemeinsame Erfahrung nicht teilen. Klumpungseffekte können als Effekt der Eingangsselektivität für eine Gruppe aber auch – wie skizziert – durch gemeinsame (Lern-)erfahrungen entstehen. Als empirisches Maß des Klumpungseffekts wird die *Intra-Class-Correlation* (ICC) herangezogen. Sie kann Werte von 0 (es gibt keine Klumpungseffekte) bis 1 (es gibt keine Unterschiede zwischen den Mitgliedern einer Gruppe, alle Unterschiede sind nur zwischen den Gruppen zu beobachten) annehmen. In Bezug auf Kompetenzen von Schülerinnen und Schülern in Deutschland sind ICCs von ca. 0,12 bis 0,20 erwartbare Größen.

Im Rahmen der bei den Schulinspektionen ermittelten Informationen z. B. zur Zufriedenheit der Lehrkräfte und auch der durch die Lehrkräfte wahrgenommenen Führung durch die Schulleitung sind ebenfalls bedeutende *Intra-Class-Correlations* zu erwarten. Diese können durch die Zusammensetzung von Kollegien z. B. hinsichtlich ihrer Ausbildung (Studienort und Referendariat) oder auch ihrer Motivation, an eine bestimmte Schule zu gehen, entstehen. Gerade Letzteres ist ein Effekt, der mit der Veröffentlichung von Ergebnissen von Leistungsdaten der Schulen in Zusammenhang zu bringen ist. Während Eltern insbesondere aus bildungsfernen Familien kaum die Möglichkeit sehen bzw. nutzen, ihre Kinder an spezifisch ausgewählten Schulen anzumelden, ist die Entscheidung der Lehrkräfte, an eine bestimmte Schule zu gehen, durchaus einerseits von (typischen) Motiven und Interessen der Lehrkräfte und andererseits durch die zunehmende Veröffentlichung von Daten über die Schulen beeinflusst (vgl. für die Vereinigten Staaten von Amerika z. B. Frankenberg 2006).

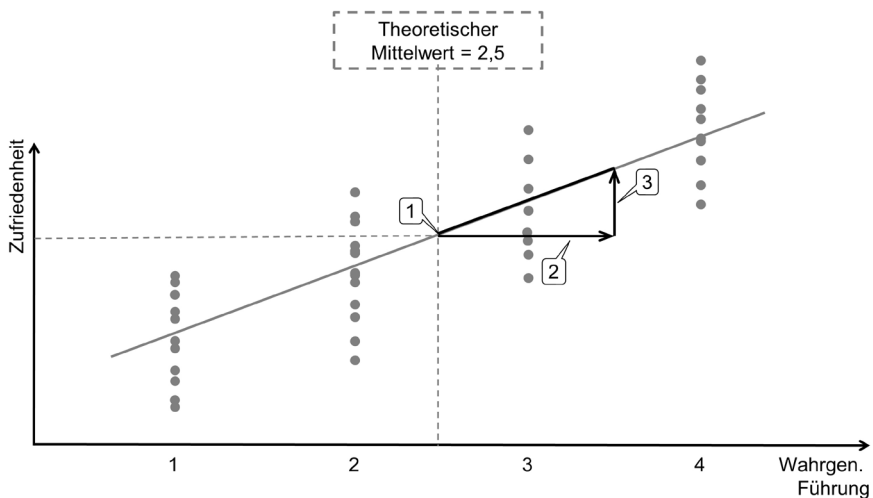
1.3 Die Methode der Regressionsanalyse

Um Hypothesen über den Zusammenhang von zwei Merkmalen zu überprüfen, bieten sich bei empirischen Untersuchungen verschiedene Methoden an. Eine sehr oft angewandte Methode ist die Regressionsanalyse, wenn Aussagen über den Einfluss einzelner oder mehrerer Variablen auf ein betrachtetes Kriterium (=eine abhängige Variable) untersucht werden sollen.

In Abbildung 3 ist das Prinzip der Regressionsanalyse abgebildet. Von befragten Lehrkräften liegen in diesem Beispiel jeweils paarweise Informationen über ihre berufliche Zufriedenheit und die von ihnen wahrgenommene Führung

der Schulleitung vor. Ziel des Regressionsverfahrens ist es, anhand eines Modells (lineares Gleichungssystem) die Unterschiede (Varianz) einer zu beschreibenden Variable (hier Zufriedenheit der Lehrkräfte) anhand einer anderen Variable (hier Führungshandeln der Schulleitungen) vorherzusagen. Die Regressionssteigung gibt dabei ein Maß der Abhängigkeit der Zufriedenheit der Lehrkräfte vom durch die Lehrkräfte wahrgenommenen Führungshandeln der Schulleitungen wieder. Beim theoretischen Skalenmittelwert von 2,5 der wahrgenommenen Führung gibt die Regressionsgerade (s. Abbildung 3) an der gestrichelten Linie ① den Grad der damit im Zusammenhang stehenden Zufriedenheit der Lehrkräfte wieder. Nimmt der Grad der wahrgenommenen Führung um eine Skaleneinheit auf 3,5 zu ②, verändert sich die Zufriedenheit in diesem Beispiel positiv ③. In den folgenden Abbildungen, die anhand der realen Inspektionsdaten berechnet wurden, wird nunmehr nur der in Abbildung 3 dunkel dargestellte Teil der Regressionsgeraden wiedergegeben, da dieser alle relevanten Informationen über den Zusammenhang der von den Lehrkräften wahrgenommenen Führung der Schulleitungen und ihrer beruflichen Zufriedenheit enthält.

Abbildung 3: Visualisierung der Analysemethode „Regressionsanalyse“



In Abhängigkeit davon, wie eine zu untersuchende Stichprobe bei den Analysen gruppiert wird, lassen sich unterschiedliche Effekte beobachten. In diesem Zusammenhang verweist Achtenhagen (1981, S. 321) auf eine Studie von Robinson, in der dieser zeigt, wie es bei Untersuchungen aufgrund von verschiedenen Gruppierungen einer Stichprobe dazu kommt, „daß der Korrelationskoeffizient zwischen

0,20 und 0,95 schwankt, je nachdem ob die Daten für Individuen, einzelne Staaten oder Gruppierungen von Staaten aufbereitet werden“. Bei der Bewertung, welcher Befund richtig ist, kommt es zunächst auf die im Fokus stehende Frage an, in der sich unterschiedliche Voraussetzungen bzw. Perspektiven widerspiegeln. Wichtig ist, dass die jeweils zur Forschungsfrage passende Analysemethode ausgewählt wird. Die Entscheidung für oder wider eine dieser Methoden – und darauf sei hier ausdrücklich hingewiesen – ist nicht beliebig. Die Auswahl folgt einer Abwägung von fachlich und methodisch zu berücksichtigenden Faktoren.

Augenfällig wird die besondere Datenstruktur, wenn neben den Lehrkräften auch die Schulleitungen in den Blick genommen werden sollen. Bei der Zusammenführung bestehen im Wesentlichen zwei Möglichkeiten: Zum einen können die Daten der Lehrkräfte je Schule gemittelt und mit den Daten der Schulleitungen zusammengespielt werden. Oder es können die Daten der Schulleitungen jeweils an die Daten der Lehrerinnen und Lehrer kopiert werden. Mit diesen Vorgehensweisen sind jeweils wesentliche Nachteile verbunden: Eine Zusammenfassung der Eigenschaften von Lehrkräften oder Merkmalen innerhalb einer Schule zu einem einzigen Wert zieht bedeutende Datenverluste nach sich, da alle individuellen Unterschiede zwischen den Lehrkräften innerhalb der Schulen verloren gehen. Zudem wird in den anschließenden Analysen nur mit ‚der durchschnittlichen Lehrkraft‘ gerechnet. Bei der Vervielfältigung von Schulleitungs- bzw. Strukturmerkmalen der Schule muss implizit vorausgesetzt werden, dass diese Merkmale auf jede Lehrkraft gleichermaßen wirken. In Bezug auf die Lehrkraft-Schulleitungs-Interaktion würde dies bedeuten, dass das von der Schulleitung gezeigte Verhalten auf jede Lehrkraft in der Schule den gleichen Einfluss ausübt. Eine Voraussetzung, die bisherigen Erfahrungen widerspricht. Und schließlich kommt es durch die Vervielfältigung der Schulleitungsdaten systematisch zu fehlerhaften Signifikanzschätzungen bei nachfolgenden Analysen.

Die in der Literatur übereinstimmende Forderung ist die, dass bei der Untersuchung von (hierarchisch) geschachtelten Daten deutlich zwischen den Ebenen zu unterscheiden ist. Bei bestimmten Fragestellungen ist danach zu berücksichtigen, dass z. B. Lehrkräfte in Schulen oder Schulen in Schulformen etc. organisiert sind. Unterbleibt eine solche analytische Trennung, kann es bei der Analyse z. B. von Zusammenhangshypothesen zu Fehlinterpretationen kommen.

Die hier vorgestellten Analysen der Schulinspektion lenken schrittweise den Blick auf verschiedene Ebenen der Schulinspektion. Als anspruchsvolles Standardverfahren bei der Analyse von Schulvergleichsuntersuchungen hat sich die sogenannte Mehrebenenanalyse bewährt, auf deren Darstellung hier jedoch verzichtet wird. Ein Anwendungsfall im Rahmen der Schulforschung ist bei Schwippert (2008) beschrieben.

1.4 Die verwendeten Skalen

Um die Aufmerksamkeit auf einen relevanten Aspekt im Hamburger Schulsystem zu lenken, wird auf das oben vorgestellte Modell des Einflusses von Schulleitungen auf den Lernerfolg von Schülerinnen und Schüler zurückgegriffen, in dem der Einfluss des Führungshandelns von Schulleitungen auf die Motivation von Lehrkräften eine wichtige Rolle spielt. Exemplarisch wird somit aus den Daten der Hamburger Schulinspektion die Skala des Schulleitungshandelns in Fragen der Führung und die Skala zur Zufriedenheit der Lehrkräfte (aus der sich Einflüsse auf motivationale Merkmale ableiten lassen) genutzt.

Das Führungshandeln von Schulleitungen wird anhand einer Reihe von Fragen, die diesbezüglich an Lehrkräfte gestellt wurden, operationalisiert und in eine Skala zusammengefasst.¹ Tabelle 3 gibt Auskunft über die in die Skala zum Führungshandeln eingehenden Items. Alle Fragen konnten auf einer vierstufigen Skala beantwortet werden, wobei der niedrigste Wert für Ablehnung bzw. Nicht-Zustimmung und der höchste für Akzeptanz bzw. Zustimmung stehen. Die über Mittelwertanalysen ermittelten Werte sind entsprechend zu interpretieren. Die Skala „Wahrgenommene Führung“ weist eine Reliabilität von $r_{tt(\text{Führung})} = 0,846$ (11 Items) auf.

Tabelle 3: Variablen der Skala „Wahrgenommene Führung im Schulleitungshandeln“

Variablenname	Frageformulierung im Lehrkräftefragebogen
LF010340	Die Schulleiterin/der Schulleiter sorgt für einen guten Informationsfluss an der Schule.
LF010350	Die Schulleiterin/der Schulleiter hat den Überblick über das Geschehen an der Schule.
LF010360	Die Schulleiterin/der Schulleiter vertritt ein klares pädagogisches Konzept.
LF010380	Die Schulleiterin/der Schulleiter repräsentiert die Schule gut nach außen.
LF010400	Die Schulleiterin/der Schulleiter ist bei wichtigen Fragen ansprechbar.
LF010021	An unserer Schule haben wir Entwicklungsprioritäten erarbeitet.
LF010441	An unserer Schule werden die Lehrerinnen und Lehrer bei wichtigen Entscheidungen systematisch beteiligt.
LF010131	An unserer Schule sind Zuständigkeiten für Entwicklungsmaßnahmen klar festgelegt.
LF010030	An unserer Schule gibt es ein systematisches Qualitätsmanagement, für das klar geregelt ist, wer was macht.
LF010230	Ich bin der Meinung, dass wir an unserer Schule momentan die zentralen Ziele und Probleme angehen.
LF010250u	Ich bin der Meinung, dass die Planung von Innovationen an unserer Schule ziemlich unsystematisch geschieht (umgepoltes Item).

1 Die Fragebögen sind im Internet einsehbar: www.schulinspektion.hamburg.de/index.php/article/detail/1006 [11.8.2014]

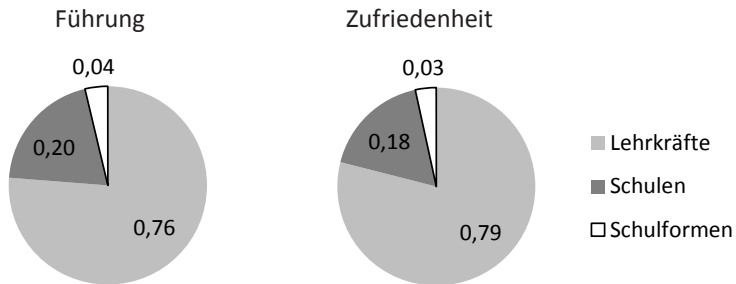
Tabelle 4: Variablen der Skala „Zufriedenheit der Lehrkräfte“

Variablenname	Frageformulierung im Lehrkräftefragebogen
LF010310	Ich bin der Meinung, dass die Räumlichkeiten der Schule und ihr Gelände einladend und freundlich sind.
LF011010	Ich bin zufrieden mit den allgemeinen Arbeitsbedingungen an der Schule.
LF011020	Ich bin zufrieden mit dem Arbeitsklima an meiner Schule.
LF011030	Ich bin zufrieden mit den Kommunikationsstrukturen an der Schule.
LF011040	Ich bin zufrieden mit meinen Möglichkeiten, Ideen zur Entwicklung der Schule einzubringen.
LF011050	Ich bin zufrieden mit meinen Möglichkeiten, berufliche Ziele, die sich auf Schule und Unterricht beziehen, verwirklichen zu können.
LF011060	Ich bin zufrieden mit der fachlichen Unterstützung durch die Kolleginnen und Kollegen.
LF011080	Ich bin zufrieden mit der technischen Ausstattung der Fach- und Unterrichtsräume.
LF011090	Ich arbeite gern an dieser Schule.

Auch die Fragen aus Tabelle 4 konnten auf einer vierstufigen Skala beantwortet werden, wobei der niedrigste Wert wieder für Ablehnung bzw. Nicht-Zustimmung und der höchste für Akzeptanz bzw. Zustimmung stehen. Die über Mittelwertanalysen ermittelten Werte sind entsprechend zu interpretieren. Die Skala zur Zufriedenheit der Lehrkräfte weist eine Reliabilität von $r_{tt(\text{Zufriedenheit})} = 0,859$ (9 Items) auf.

Die Fragen, ob sich die Kollegien in der Wahrnehmung der Führung durch die Schulleitung und in Bezug auf die Zufriedenheit der Lehrkräfte systematisch zwischen den Schulen und darüber hinaus auch noch zwischen den Schulformen unterscheiden, lässt sich anhand der oben vorgestellten *Intra-Class-Correlation* bestimmen. Hier zeigt sich, dass – neben den individuellen Unterschieden in der Wahrnehmung des Führungshandeln und der Zufriedenheit innerhalb der Schulen – auch ein substanzieller Anteil systematisch (signifikant) zwischen den Schulen zu beobachten ist ($ICC_{\text{Führung-Schule}} = 0,20 / ICC_{\text{Zufriedenheit-Schule}} = 0,18$). Wenn auch geringer, so lassen sich auch systematische (signifikante) Effekte zwischen den in der Hamburger Schulinspektion berücksichtigten Schulformen ($ICC_{\text{Führung-Schulform}} = 0,04 / ICC_{\text{Zufriedenheit-Schulform}} = 0,03$) feststellen. Die auf den verschiedenen Ebenen zu beobachtenden Varianzanteile der durch die Lehrkräfte wahrgenommenen Führung und der Zufriedenheit der Lehrkräfte sind Abbildung 4 zu entnehmen.

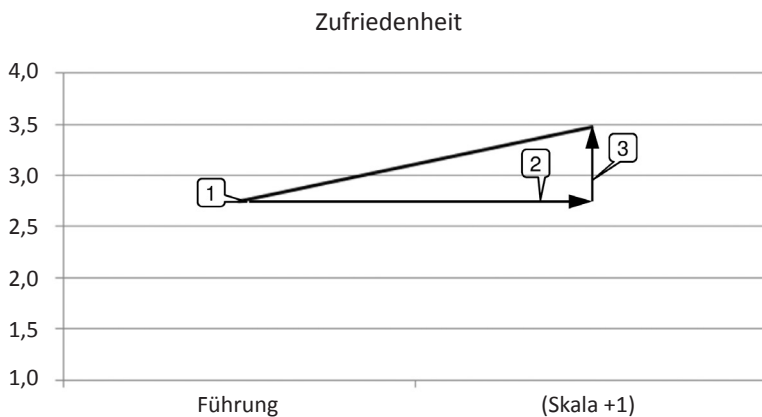
Abbildung 4: Varianzanteile (ICCs) der wahrgenommenen Führung und der Zufriedenheit der Lehrkräfte innerhalb von Schulen, zwischen Schulen und zwischen Schulformen



2. Befunde

In Abbildung 5 ist der generelle Zusammenhang zwischen der von den Lehrkräften wahrgenommenen Führung (Abszisse) und ihrer Zufriedenheit (Ordinate) visualisiert.

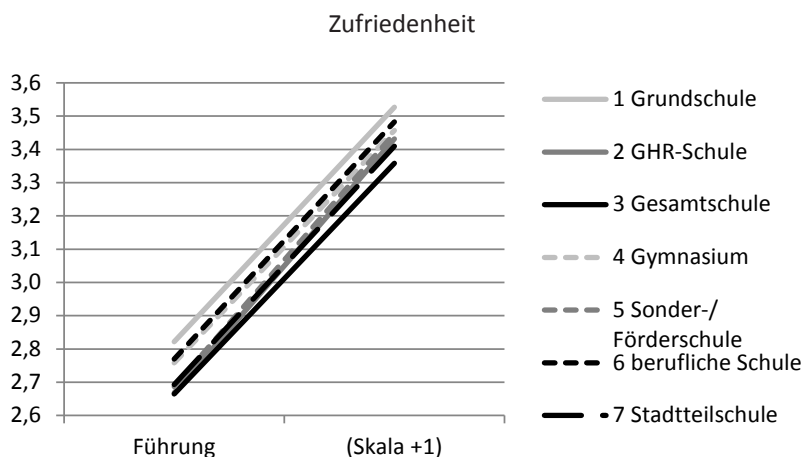
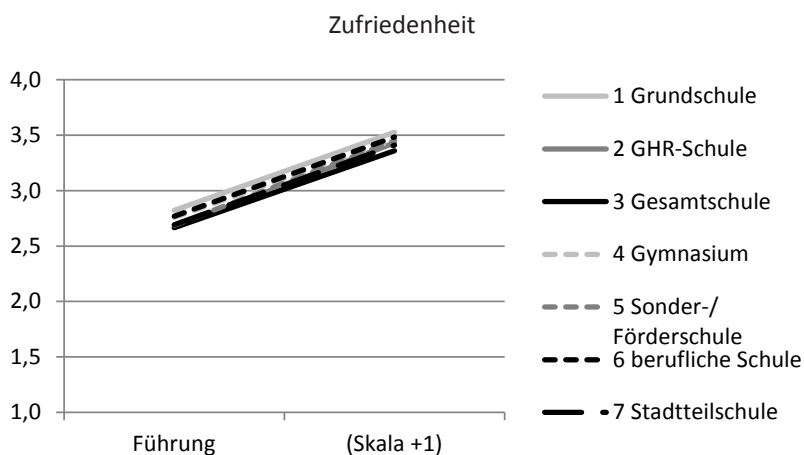
Abbildung 5: Zusammenhang von wahrgenommenem Führungshandeln und Zufriedenheit von Lehrkräften



Bei Lehrkräften, die auf der Skala für die wahrgenommene Führung einen theoretischen Skalenmittelwert aufweisen (2,5), liegt der erwartete Wert für ihre Zufriedenheit bei 2,75 (vgl. ①) und damit etwas über dem erwarteten Skalenmittelwert.

Für Lehrkräfte, für die ein um 1 höherer Skalenmittelwert von 3,5 (vgl. ②) für die Führung ermittelt wird, liegt die erwartete Zufriedenheit mit der wahrgenommenen Führung um 0,72 höher (vgl. ③) – und somit bei insgesamt 3,47 Skalenpunkten. Der im Mittel von den Lehrkräften angegebene Wert für die wahrgenommene Führung liegt bei 3,00 und damit ebenfalls leicht über dem erwarteten, theoretischen Skalenmittelwert – was auf eine tendenziell positive Einschätzung der Führung durch die Schulleitungen hindeutet.

Abbildung 6: Zusammenhang von wahrgenommenem Führungshandeln und Zufriedenheit von Lehrkräften nach Schulformen (Originalskala und Skalenausschnitt)



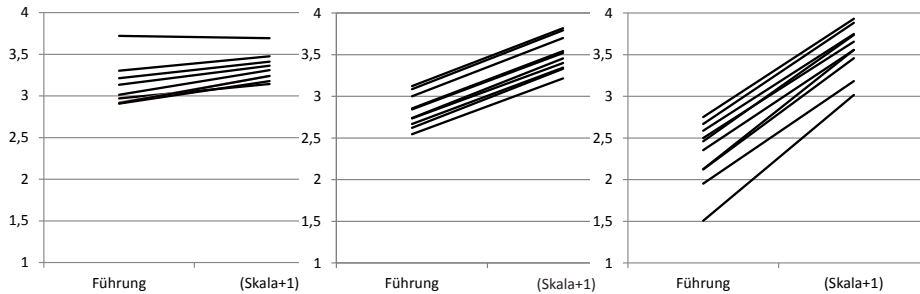
Der Abbildung 5 ist zu entnehmen, dass mit zunehmend positiver Wahrnehmung von Führung durch die Lehrkräfte auch die Zufriedenheit der Lehrkräfte an der Schule zunimmt.

Der generelle Zusammenhang von Führung und Zufriedenheit kann auch auf kleinere Betrachtungseinheiten (Schulformen) heruntergebrochen werden. Diese Differenzierung ist in Abbildung 6 dargestellt. Jede der kleineren Linien stellt hier eine Untergruppe (Schulform) der gesamten dargestellten Daten (hier aus der Hamburger Schulinspektion) dar. Bei diesen und den nachfolgenden Analysen wurden von den 163 Grundschulen 22 Schulen ausgeklammert, wenn weniger als 10 Lehrkräfte den Fragebogen pro Schule beantwortet haben, da es hierbei durch einzelne extreme Antworten zu verzerrten Schätzungen des Schuleffektes kommen kann. In Bezug auf die Lehrkräfte bedeutet dies, dass von den 2.592 befragten Grundschullehrkräften 162 nicht berücksichtigt werden konnten.

In Abbildung 6 (oben) sind wiederum die für die einzelnen Gruppen bestimmten Regressionsgeraden eingezeichnet (siehe den unteren Teil der Abbildung 6, in der die Steigungen vergrößert dargestellt sind). In dem hier gewählten Beispiel verlaufen alle Regressionsgeraden für jede der Schulformen von unten links nach oben rechts und zeigen damit einen (rechnerisch) positiven Zusammenhang zwischen wahrgenommenem Führungshandeln und Zufriedenheit an, auch wenn sich diese zwischen den verschiedenen Schulformen leicht unterscheidet. Je nach Schulform nimmt die Zufriedenheit der Lehrkräfte zwischen 0,69 und 0,76 Punkten zu, wenn sich im Kollegium die wahrgenommene Führung um einen Skalenpunkt erhöht.

Um einen Eindruck zu vermitteln, wie stark sich der Zusammenhang zwischen der wahrgenommenen Führung und der Zufriedenheit der Lehrkräfte jeweils zwischen den Einzelschulen unterscheidet, wurden in Abbildung 7 jeweils die 10 Schulen dargestellt, die den niedrigsten Zusammenhang aufweisen (flacher Verlauf: von $-0,03$ bis $0,34$), die einen mittleren Zusammenhang aufweisen ($0,67$ bis $0,72$) und schließlich die, die den höchsten Zusammenhang (steilster Verlauf: von $1,16$ bis $1,51$) aufweisen.

Abbildung 7: Zusammenhang von wahrgenommenem Führungshandeln und Zufriedenheit von Lehrkräften nach Grundschulen in Gruppen mit niedrigem, mittlerem, starkem Zusammenhang



Wird der globale Eindruck des Zusammenhangs (Abbildung 5) von wahrgenommener Führung und Zufriedenheit den Zusammenhängen je Schulform (Abbildung 6) und schließlich je Grundschule (Abbildung 7) gegenübergestellt, wird deutlich, wie wichtig es ist, verschiedene Analyseebenen zu differenzieren. Denn obwohl die Aussage einmal lautet, der Zusammenhang von wahrgenommenem Führungshandeln und Zufriedenheit ist positiv, lautet sie bei anderer Betrachtung, dass es durchaus Bildungseinrichtungen gibt, in denen dieser Zusammenhang nicht so (deutlich) ausgeprägt ist. Trotz dieses Widerspruchs: Beide Aussagen sind richtig – sie beziehen sich nur auf unterschiedliche Ebenen des Bildungssystems und haben damit auch unterschiedliche Reichweiten bei der Interpretation.

3. Zusammenfassende Betrachtung

Die vorgestellten Befunde zum Zusammenhang der von den Lehrkräften wahrgenommenen Führung durch die Schulleitung und ihrer Zufriedenheit lassen sich analog den vorgestellten Bezugsnormorientierungen interpretieren. Wird ein globaler Zusammenhang (über alle Schulen und Lehrkräfte hinweg) ermittelt, stellt sich die normative Frage, wie stark die Zufriedenheit der Lehrkräfte mit dem Führungshandeln der Schulleitungen zusammenhängen darf oder soll. Ist ein ausgeprägter Zusammenhang gut oder ist er eher nachteilig? Was bedeutet der festgestellte Zusammenhang für die Aus- und Weiterbildung von Schulleitungen in Bezug auf Fragen der Führung? Und noch globaler gefragt: Welchen Stellenwert nimmt die Zufriedenheit der Lehrkräfte im gesamten Kanon von Zielgrößen innerhalb der Schule ein? Ist die Zufriedenheit bereits Selbstzweck oder Mittel zum Zweck? Das Modell von Leithwood zeigt, dass die Zufriedenheit der Lehrkräfte sich, vermittelt

über ihr unterrichtliches Handeln, auf die Entwicklung der Schülerinnen und Schüler auswirkt – und somit auf *das* zentrale Ziel für alle Akteure im Schulsystem zielt.

Der Vergleich zwischen den Schulformen zeigt, dass es kaum Unterschiede zwischen den Schulformen in der Stärke des Zusammenhangs zwischen Schulleitungshandeln und Zufriedenheit der Lehrkräfte gibt; auch das Durchschnittsniveau der Zufriedenheit der Lehrkräfte unterscheidet sich zwischen den Schulformen nur um maximal 0,28 Punkte (also um rund einen Viertel-Skalenpunkt).

Lenkt man den Blick auf die individuelle Bezugsnorm, rückt im Analogieschluss die Einzelschule in das Zentrum der Betrachtung. Bei dieser Betrachtung bekommen die innerschulischen Differenzen im Zusammenhang zwischen Führung und Zufriedenheit eine ganz andere Bedeutung. Zum einen ist bei der Betrachtung zu hinterfragen, warum die Wahrnehmung der von der Schulleitung praktizierten Führung durch die Lehrkräfte sich im Kollegium (deutlich) unterscheidet und warum sich die wahrgenommene Führung in einigen Schulen stark auf die Zufriedenheit der Lehrkräfte auswirkt und in anderen (gar) nicht. Hier helfen Vergleiche (im Sinne einer sozialen Bezugsnorm) mit anderen Schulen, da bei der Erklärung dieser Merkmale eine Reihe von weiteren Merkmalen betrachtet werden muss, die bei dieser Analyse (noch) nicht berücksichtigt sind.

Die zentrale Erkenntnis dieses Beitrags ist, dass es keine bessere oder schlechtere Methode bei der Analyse von Daten aus den Schulinspektionen gibt. Vielmehr ist darauf zu achten, dass die der Fragestellung angemessene Methode auszuwählen ist. Hierbei ist, wie im vorgestellten Datenbeispiel abschließend angemerkt, nicht nur die eigentliche Fragestellung für sich in den Blick zu nehmen, sondern auch die Strukturmerkmale, die potenziell einen Einfluss auf den analysierten Befund nehmen können bzw. mit diesem in einem inhaltlichen Zusammenhang stehen.

Welche Implikationen ergeben sich hierdurch für die zukünftige Erstellung von Schulinspektionsberichten und den zukünftigen Umgang mit Schulinspektionen?

Abhängig von der Zielvariablen muss im Vorwege überlegt werden, ob die beobachteten Merkmale an einer normativen (kriterialen) Vorgabe gemessen werden, ob sie dazu herangezogen werden sollen, im Vergleich mit anderen Bildungseinrichtungen relative Stärken und auch Schwächen zu erkennen, oder ob es schließlich darum gehen soll, innerschulische Zusammenhänge zu analysieren. Die Frage, welche Analysemethode für die jeweils im Fokus stehende Fragestellung angemessen ist, gilt unabhängig von der Bezugsnormorientierung sowohl für eine Inspektions-interne Auswertung als auch für Auswertungen im Rahmen von (wissenschaftlichen) Re-Analysen der Hamburger Schulinspektionsdaten.

Kriteriale bzw. Normvorstellungen für das Bildungssystem sind gesellschaftlich geprägt. Möchte man an dieser Stelle analytisch bei der Schulinspektion anset-

zen, sind konsensual Ziele bzw. (z. B. Verhaltens-)Standards zu setzen, die dann als Zielgrößen für das gesamte Bildungssystem anzuerkennen sind. Bei dieser Form der Auseinandersetzung mit dem Bildungssystem spielen die für das Hamburger Schulsystem verfassten Bildungsberichte eine wichtige Rolle, da sie zum einen Bestandsanalysen vorlegen, aber auch Entwicklungen, die durch gesellschaftliche Erwartungen und auch durch sich wandelnde Voraussetzungen der verschiedenen Akteure im Bildungssystem beeinflusst sind, aufgreifen und somit einer konstruktiven Diskussion zuführen. Darüber hinaus könnten Rückmeldungen an Schulen genutzt werden, um im Vergleich mit anderen Schulen eine Bestandsanalyse zu machen. Hierbei ist es für Schulleitungen, Lehrkräfte, Eltern, Schülerinnen und Schüler hilfreich, Informationen der eigenen Schule im Vergleich mit anderen (ähnlichen) Schulen zu erhalten. So könnten z. B. anhand der Analyse von „erfolgreichen“ Schulen Hinweise auf gute Praxis herausgearbeitet werden, die ggf. für die Entwicklung der eigenen Schule genutzt werden können. Intellektuell problematisch, wenn auch analytisch sinnvoll ist es, auch einen vergleichenden Blick auf die Schulen zu werfen, die nicht erfolgreich sind. Problematisch wird deren Identifizierung jedoch, wenn damit eine Stigmatisierung einhergeht. Dennoch erscheint die Identifizierung auch „problematischer“ Schulen notwendig, um überhaupt die Chance zu haben, diesen Unterstützung seitens anderer Schulen oder der Behörde zukommen zu lassen. Auch wenn die Identifizierung nicht erfolgreicher Schulen für einzelne Personen oder Personengruppen belastend sein dürfte, so ist doch das gesellschaftlich vereinbarte Ziel, dass es zentral um das Wohl und die positive Entwicklung von Schülerinnen und Schülern geht, in den Vordergrund zu stellen. Gerade im Zusammenhang mit solchen Schulen erscheint es sinnvoll, diese zu identifizieren und mit Unterstützung durch geschulte Moderatoren den Beratungsprozess begleiten zu lassen, damit (komplexe) statistisch ermittelte Daten nicht falsch oder überinterpretiert werden.

Von einer Rückmeldung oder Besprechung von individuellen Antworten im schulischen Kontext ist abzuraten. Durch das Wissen, dass die im Rahmen der Inspektion gestellten Fragen individuell ausgewertet werden, kann es zur Orientierung an sozialen Normen und damit zur Beantwortung der Fragen gemäß der sozialen Erwünschtheit kommen. Dies kann nicht im Interesse der Schulinspektionen sein. Auch ohne die Anonymität aufzuheben, kann der Fundus an Informationen, der in den Einzelinformationen liegt, genutzt werden. Hierbei können geschulte Moderatoren helfen, die Heterogenität im Kollegium zu analysieren, ohne die Anonymität der Einzelnen aufzuheben. So könnten wichtige Informationen (z. B.: Warum wird Führung durch die Schulleitung so unterschiedlich wahrgenommen?) genutzt werden, um aktive Schulentwicklung zu betreiben.

Ein weites Forschungsfeld öffnet sich, wenn der Frage nachgegangen wird, wie Schulleitungen auf das Schulleitungshandeln vorzubereiten sind und – noch globaler angesetzt – was überhaupt Erwartungen an „gutes“ Schulleitungshandeln

ausmacht. Der Blick auf internationale Schulvergleichsuntersuchungen zeigt, dass es hier durchaus ganz verschiedene Vorstellungen betreffend des Schulleitungshandelns gibt und dass mit diesen unmittelbar auch die jeweilige Aus- bzw. Weiterbildung verknüpft ist. Hier reicht die Bandbreite von Lehrkräften mit erweiterten Verwaltungsaufgaben bis hin zu Managern von Bildungseinrichtungen (vgl. z. B. Lankes & Carstensen 2007).

Ein bisher durch Schul- und Unterrichtsforschung wenig systematisch untersuchtes Forschungsdesiderat besteht in der Auseinandersetzung mit der Frage, was die Erwartungen der verschiedenen Akteure im Bildungssystem in Bezug auf die Inspektion sind. Diese Frage erscheint für die Schulinspektionen substanziell, da sie den Blick auf die wichtigen und vor allem richtigen Stellen im Bildungssystem lenkt. Neben sowohl normativ als auch bildungspolitisch zu führenden Diskussionen bieten die bisher erhobenen Daten aus Schulinspektionen einen Fundus für eine ganze Reihe von Antworten – wobei die Herausforderung darin besteht, die richtigen Fragen zu stellen.

Literatur

- Achtenhagen, F. (1981). Mehrebenenanalyse in der Unterrichtsforschung. *Unterrichtswissenschaft*, 4, 319–336.
- Adams, R. J. & Wu, M. L. (2000). *PISA 2000 Technical Report*. Verfügbar unter: www.pisa.oecd.org [09.08.2014]
- Behörde für Bildung und Sport (2006). *Orientierungsrahmen – Qualitätsentwicklung an Hamburger Schulen*. Hamburg: Behörde für Bildung und Sport.
- Cochran, W. G. (1972). *Stichprobenverfahren*. Berlin: de Gruyter.
- Cronbach, L. J. (1976). *Research on classrooms and schools. Formulation of questions, design, and analysis*. Stanford: (Occasional Paper).
- Ditton, H. (1998). *Mehrebenenanalyse*. Weinheim: Juventa.
- Frankenberg, E. (2006). *The Segregation of American Teachers*. Verfügbar unter: <http://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/the-segregation-of-american-teachers/frankenbergorfield-segregation-american-teachers.pdf> [09.08.2014]
- Institut für Bildungsmonitoring (2009). *Jahresbericht der Schulinspektion, Hamburg 2008*. Hamburg: Behörde für Schule und Berufsbildung.
- Institut für Bildungsmonitoring (2011). *Jahresbericht der Schulinspektion, Hamburg 2009–2010*. Hamburg: Behörde für Schule und Berufsbildung.
- Institut für Bildungsmonitoring und Qualitätsentwicklung (2012). *Jahresbericht der Schulinspektion, Hamburg 2010–2011*. Hamburg: Behörde für Schule und Berufsbildung.
- Janke, N. (2006). *Soziales Klima an Schulen aus Lehrer-, Schulleiter- und Schülerperspektive*. Münster: Waxmann.

- Kish, L. (1995). *Survey sampling*. New York: John Wiley & Sons (Nachdruck).
- Krapp, A. & Weidenmann, B. (Hrsg.) (2001). *Pädagogische Psychologie* (4. Aufl.). Weinheim: Beltz PVU.
- Lankes, E.-M. & Carstensen, C. H. (2007). Der Leseunterricht aus der Sicht der Lehrkräfte. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 161–193). Münster: Waxmann.
- Martin, M. O., Mullis, I. V. S. & Kennedy, A. M. (2007). *PIRLS 2006 Technical Report*. Chestnut Hill, MA: Boston College.
- Martin, M. O., Rust, K. & Adams, R. J. (Hrsg.) (1999). *Technical standards for IEA studies*. Amsterdam: International Association for the Evaluation of Educational Achievement.
- Pietsch, M. (2014). Wirksame Schulleitungen. Annahmen, Befunde und der Blick der Schulinspektion. *Hamburg macht Schule*, 1 (26), 6–11.
- Rheinberg, F. (2001). Bezugsnormen und schulische Leistungsbeurteilung. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 59–71). Weinheim: Beltz.
- Schwippert, K. (2002). *Optimalklassen: Mehrebenenanalytische Untersuchungen*. Münster: Waxmann.
- Schwippert, K. (2008). Interpretationen und Fehlinterpretationen von Daten: Mehrebenenanalysen in der Schul- und Unterrichtsforschung. In F. Hellmich (Hrsg.), *Lehr-Lernforschung und Grundschulpädagogik* (S. 157–167). Bad Heilbrunn: Klinkhardt.
- Treiber, B. (1980). *Qualifizierung und Chancenausgleich in Schulklassen. Teil 1: Theorien, Methoden, Ergebnisse*. Frankfurt a. M.: Lang.